

EXHIBIT Q

9/17/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

Melanie Kambadur

Page 1

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION

Richard Kadrey, et al.,)
Individual and Representative)
Plaintiffs,) CASE NO.
) 3:23-cv-03417-VC
-against-)
)
Meta Platforms, Inc.,)
Defendant.)
)

*** HIGHLY CONFIDENTIAL ***

ATTORNEYS' EYES ONLY
VIDEO-RECORDED DEPOSITION OF
MELANIE KAMBADUR

Cooley, LLP
55 Hudson Yards
New York, New York 10001

09/17/2024
9:07 a.m. (EDT)

REPORTED BY: MONIQUE CABRERA

DIGITAL EVIDENCE GROUP
1730 M Street, NW, Suite 812
Washington, D.C. 20036
(202) 232-0646

9/17/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

Melanie Kambadur

Page 116

1 Q. Now, the next bullet point says
2 "Higher quality."

3 A. Yes.

4 Q. Does this mean that for Llama 3,
5 your goal was to obtain higher quality data for
6 training -- pre-training Llama 3?

7 A. What do you mean by "obtain"?

8 Q. Using your own words, to collect or
9 process.

10 A. Yes, to collect or process.

11 Q. What do you mean by "collect or
12 process," by the way?

13 A. I mean that some of this data we may
14 have already had available and just not put it
15 into the model before or processed it before.

16 Q. So, for example, there may have been
17 datasets that you were considering using for
18 pre-training that you had not incorporated yet?

19 A. Yes.

20 Q. Or processed it. Excuse me.

21 A. Either one.

22 Q. For example, Libgen, is that a

9/17/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

Melanie Kambadur

Page 117

1 dataset that Meta had, but had not processed yet?

2 A. I don't recall when we processed it.

3 Q. Now, do you recall testifying
4 earlier that Libgen was used to pre-train
5 Llama 3, Llama 3 series of models?

6 A. Yes. At least a subset of Libgen
7 was used.

8 Q. But it wasn't used to pre-train the
9 Llama 2 models or the Llama 1 models, right,
10 according to your testimony today?

11 A. Yes, that's correct, to my
12 knowledge.

13 Q. Now, do you remember if Meta had the
14 Libgen dataset before it was processed and
15 including in the Llama 3 models?

16 MR. WEINSTEIN: Object to form.

17 A. Of course we had to have the dataset
18 before including it in the model, yes.

19 BY MR. YOUNG:

20 Q. So that means that there was a
21 decision made to incorporate Libgen, process
22 Libgen as pre-training date for Llama 3, correct?

9/17/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

Melanie Kambadur

Page 118

1 A. Yes.

2 Q. Do you know who made that decision?

3 A. I don't think there was a

4 singular -- there was not a singular person.

5 Q. But a decision was made, correct?

6 A. To include Libgen into Llama 3?

7 Yes.

8 Q. Do you remember why that decision
9 was made to include Libgen into Llama 3?10 A. As with all of our datasets, we had
11 run some small-scale experiments that indicated
12 it would positively improve our benchmarks on a
13 larger model. And we had gone through our legal
14 and privacy review procedures.15 Q. Did anyone express to you any
16 concern with using the Libgen data source as
17 pre-training material?18 A. I recall people asking if it was
19 approved.20 Q. Did you express any concern about
21 using the Libgen data source as pre-training
22 material?

9/17/2024

Richard Kadrey, et al. v. Meta Platforms, Inc.
Highly Confidential - Attorneys' Eyes Only

Melanie Kambadur

Page 119

1 MR. WEINSTEIN: Exclude any
2 communications you may or may not have had
3 with counsel.

4 (Reporter clarification.)

5 MR. WEINSTEIN: Exclude any
6 communications you may or may not have had
7 with counsel.

8 A. I don't recall specific
9 conversations, but it's possible I also discussed
10 making sure that we went through the approval
11 processes.

12 BY MR. YOUNG:

13 Q. Did you ever discuss with anyone who
14 was not a lawyer any concerns with using Books3
15 as pre-training material for any of the Llama
16 models?

17 A. It's poss- --

18 MR. WEINSTEIN: Again, same -- same
19 instruction. His question's limiting to
20 non-attorneys, but if you are -- in answering
21 the question, if -- you shouldn't repeat
22 anything a lawyer may have told you, even